

Part 2: Description of Work

1. Summary

Biological systems consist of elements that interact within and across levels. For example, interactions among genes and plasticity determine traits of individuals and adaptations to environments and habitats, competitive and cooperative interactions among individuals influence population dynamics, and interactions among species affect the dynamics of communities and ecosystem processes. Such systems, classically studied as simple graphs, can be explored as networks of networks to capture the interactions inside and between networks, interdependencies among hierarchical levels, and in general, the connectivity patterns and processes across the different scales of biological organization (Figure 1).

BIODIVERSITY DYNAMICS

BIODIVERSITY DECLINE

DIVERSIFICATION

Statements of the goals

We propose to study interactions within and between biological levels to decipher key interdependencies that maintain biodiversity in ecosystems. We bridge data science and theories of biodiversity on the shoulders of recent studies focusing on the entangled web of life, that is, interactions and interdependencies across genetic, phenotypic, ecological and spatial scales, represented each data set as a network. Our goal is to test the need to disentangle a multilayer network to understand biodiversity, that is, a simple graph approach is not enough to understand biodiversity. The understanding of the webs of life requires the entanglement of several layers of data-driven theory.

Milestones

i) Data animations To developing visualizations of genetic, phenotypic, ecological and spatial data to 1) Gain insights on how the the interactions connecting inside and across biological levels drive the evolution and maintenance of biodiversity, and 2) Communicate in public exhibitions the core patterns and processes governing interdependent networks across biological levels.

ii) Data mining and NoN inference. To infer patterns of interdependence within and between genetic, phenotypic, ecological, and spatial networks **THIS IS REPETITIVE to gain insights into the connectivity within and across biological levels** (Box 1).

iii) iii) Processes driving tangled webs. To Infer mechanisms using data-driven modeling scenarios coupling genetic, phenotypic, ecological, and spatial networks **THIS IS REPETITIVE to explore how connectivity patterns within and between levels alter biodiversity patterns** (Box 2).

Significance of the project for data science

The increase in availability of biological data and computational power together with new methods presents many opportunities to strengthen the interplay between ecological and evo-

lutionary processes maintaining diversity. Yet, while data science is progressing in merging patterns and processes for the existing large-scale data in many fields, the integration of patterns and processes integrating data across biological levels in ecology and evolution is still at a very incipient stage. In this project we envision three main advances for data science: i) Strengthening the feedback between data science and data-driven synthesis in ecology and evolution. We will release a data-driven package scalable to other datasets to decipher the strength of interdependencies across biological networks, the patterns and mechanisms that underlie such interdependencies and their effects on biodiversity maintenance, and ii) Communicating throughout animations in a permanent public exhibition the importance of using interdependent biological networks in merging data science, ecological and evolutionary processes for the maintenance of biodiversity.

2. Background

The study of interactions, both within and across hierarchical scales, is central to the ongoing synthesis of ecology and evolution in general [1, 2, 3], and to debates surrounding the relationship between complexity and stability in particular. Ecologists, for example, have argued for empirical patterns of positive [4], negative [5] and non-relationships [6] between the number of links and the stability of food webs (i.e., the number of links usually defined as the complexity of the food web). This debate is rooted in the mechanisms driving ecological interactions within and among species [5, 7, 8, 9, 10, 11, 12]. Analogously, evolutionary biologists have puzzled over the relationships between the complexity of gene interactions and the stability of phenotypes [13, 14, 15, 16]. Quantitative genetics theory predicts that most genetic variance in populations is additive [17], and yet accounting for gene interactions can improve predictions about the distribution and evolution of traits [18]. Experiments are also increasingly showing that gene interactions are common and that additivity can be an emergent property of underlying genetic interaction networks [19, 20, 21, 22, 23, 24].

Although the relationship between complexity and stability has been explored within hierarchical levels, such as either individuals, populations or food webs, the relationship among levels has received less attention [25, 26, 27]. Eco-evolutionary theory usually includes interactions in one hierarchical level network, therefore we do not have a good understanding of how to integrate data and theory to connect complex traits to the stability and complexity of ecological networks. It is possible that small-scale interactions at one hierarchical level might help explain large-scale patterns at another (“Hierarchical networks” in Figure 1). For example, networks of gene interactions and migration dynamics could plausibly influence trait-dependent interactions between populations. In such hierarchical systems, what is the relationship between complexity and stability?

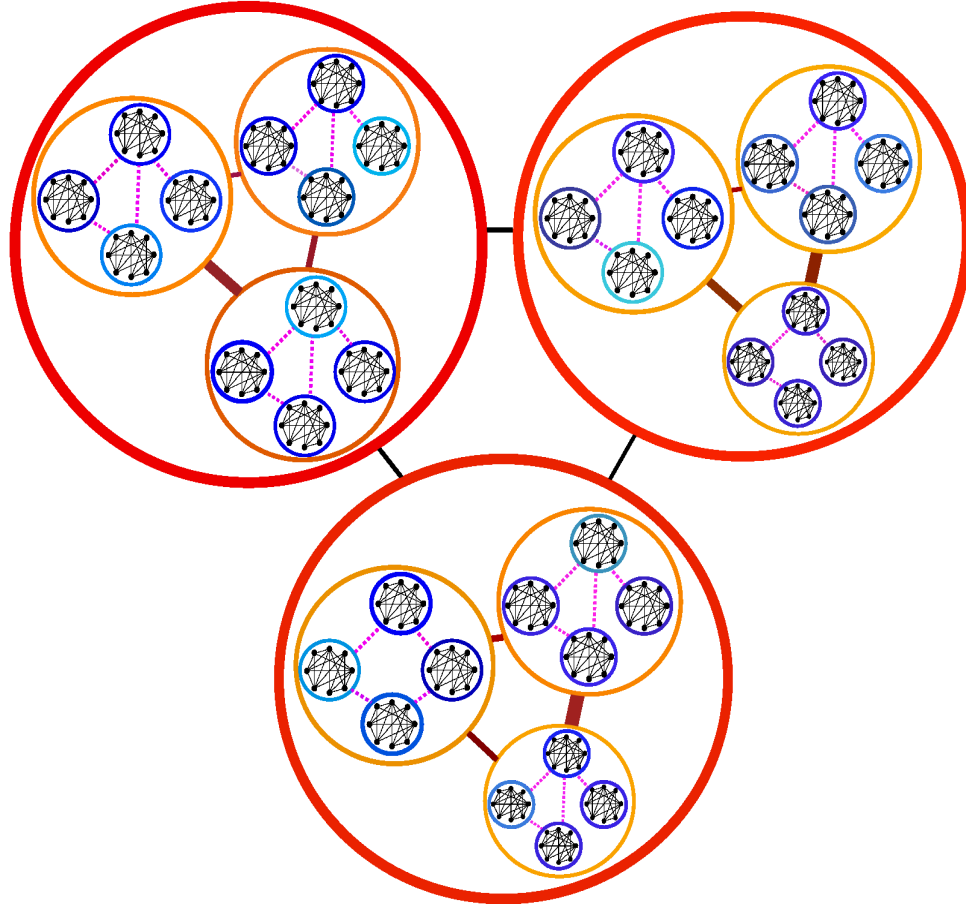


Figure 1: Hierarchical network approach: We represent a meta-ecosystem with patches (red), species (orange), individuals (blue), and genes (black). At intra-organismal level, genes interact to produce a trait (blue). At inter-organismal level individuals mate (dotted pink links) to produce the distribution of traits (blue tones). At population level trait distribution drives the interaction strength between species (brown links). The meta-ecosystem is pictured as a spatial network (black links) of local interaction networks.

3. Significance to Data Science, Scientific Goals and Objectives

Data spanning many branches of ecology and evolution are starting to decipher the interaction patterns across levels of biological organization and the relationship between complexity and stability [28, 29, 30, 31]. Many libraries are also rapidly emerging to integrate, analyze and visualize patterns across networks [28, 30, 32, 33]. For example, recent analysis of six different levels of biological organization depicting gene interactions, complex phenotypes, animal societies, metapopulations, food webs, and vertebrate communities has shown invariant patterns of nestedness independently of the interaction type or biological scale [34]. In a nested network interactions are organized such that specialists (for example predators that eat only a few prey) interact with a subset of the species with whom generalist, for example, predators that attack many preys, interact. Nestedness has received significant attention because it has been suggested that a nested pattern of interactions may lead to greater or lower biodiversity in ecological networks [35, 36]. Regardless of whether nestedness or other network pattern increase or decrease the number of species in ecological networks, it is an open question how data-driven inference methods can help to capture the interdependencies in hierarchical networks and which ecological and evolutionary processes drive such interdependency. Therefore, our the goals of this proposal are:

- Combine visualization and animations to exploratory analysis and computaton of metrics to detect network patterns crossing two or more networks using our source data hosted by *Renga Platform*.
- Team up with the SDSC staff the implementation of Markov random field methods to efficiently compute and/or reconstruct the interdependencies among genetic, phenotypic, ecological and spatial networks (Box 1).
- Implementing, debugging, and analyzing meta-ecosystem eco-evolutionary network process-based modeling scenarios parametrized using the inference patterns obtained in goal 2 (Box 2). This step will use linear algebra techniques to contrast the stability-diversity properties of multilayer networks between scenarios driven by the empirical patterns and theoretical scenarios using null or neutral models.

4. Research Plan

4.1 General description of the scientific approach

We describe our scientific approach in Box 1 highlighting 1) the integration between data-driven and process-based approaches in biodiversity research, 2) the existing model systems

to understand the multilayer nature of biodiversity, and 3) a modeling prototype bridging data-driven and process-based approaches connecting multilayer networks with deep learning networks.

Biodiversity research has been systematically studied at only one biological level and splitted in many temporal and spatial scales. This has produced an immense gain in detailed knowledge at each of the levels and scales studied, but the one-level and one-scale approach might be insufficient to understand diversification, biodiversity dynamics and biodiversity decline under human-driven disturbances. For example, we might think that the decline of one species is due to low genetic diversity in response to a rapidly changing environment. However, this decline can be triggered by the absence or the presence of specialized predators-parasites in specific habitats and by the genetic architecture of the complex traits involved in the prey-predator-parasites interactions. Taking together, to understand the decline of one species we would need not only the genetic variance, but also the community of predators-parasites and the architecture of the defense traits across the landscape. On the other hand, there are a few model systems where multidimensional individual data have been collected (See section 4.2 Data Sources, Figure 2) but datasets collected at multiple levels or scales are currently missing. This situation leaves us in a data and methods gap meaning we currently do not have datasets collected or integrating different biological levels and spatial scales to be confronted against deep process-based learning networks framing different scenarios of connections within and among layers [?].

Box 1. Deep learning and multilayer networks in Biodiversity research

Our approach will contain two steps. First, we will join data-driven inference with process-based approaches to decipher the mechanisms that best predict the patterns within and among the layers of collected data (section 4.2 Data Sources.) Second, we will run deep process-based learning models exploring a broad range of biodiversity datasets to explore the network configurations within and between layers that maintain biodiversity.

Step 1: Deep learning networks in biodiversity data We will explore networks with varying structures to infer the ones minimizing the difference with the empirical patterns DESCRIPTION OF EMPIRICAL PATTERNS. The starting point will be to explore networks allowing for feedbacks and nonlinear functions within and between layers (Figure 2). WHY THESE NETWORKS We will explore a range of topologies from bidirectional recurrent neural networks (BRNN) to feedforward neural networks (FNN) and reinforcement learning in unknown environments (RL) [?].

Inference Most data sets in ecology and evolution are large collections of small data sets (See section 4.2 Data Sources.) For example, in areas such as species ranges and species interactions, there might be a large amount of data, but there is still a relatively small amount of data for each gene, phenotype, individual or interaction. To customize predictions for species ranges and accounting for abiotic and biotic factors it becomes necessary to build scenarios accounting for the heterogeneity at individual level, with its inherent uncertainties, and to couple these models together in a hierarchy scaling from genes, to phenotypes, populations, communities and ecosystems, so that information can be borrowed from other similar levels across the landscape in the absence of empirical estimations. This individualization of models [?], will be implemented using hierarchical Bayesian neural networks approaches such as hierarchical Dirichlet processes therefore accounting for many interdependent layers.

Step 2: Deep process-based learning networks in biodiversity research We will explore a suite of deep process-based learning networks to infer the role of feedbacks between evolutionary and ecological networks on biodiversity dynamics. We will account for demography, trait evolution, gene flow and selection regimes across many interacting species across landscape gradients (Box 2). We will integrate datasets from many taxa to decipher the most plausible network configurations within and between layers that maintain biodiversity (<https://github.com/melian009/Robhoot/blob/master/layers/data.integration/databases.md>). Box 2 describes the starting point to explore further additional network topologies considering BRNN, FNN and RL type networks. We will explore how disturbance affect biodiversity loss in two different scenarios (i.e., independent layers and feedbacks). As a consequence of this exploration, we aim to advance which data will be needed to collect or integrate at different levels and scales to better understand biodiversity dynamics. Which mechanisms within and between layers are needed to predict biodiversity dynamics and its current decline? This step will help to design field studies to collect data at different levels simultaneously, from genome NGS and gene expression patterns to phenomes and individual interactions across many sites in the landscape.

Box 2. Deep process-based learning in Biodiversity research.

To explore the effect of evolutionary and ecological processes on multilayer networks, we propose to join deep learning with process-based networks taking into account demography, trait evolution, gene flow and selection. 1) gene interaction networks to a trait distribution and 2) trait distributions to interaction strength between species (Figure 3). The meta-ecosystem contains \mathcal{P} patches and \mathcal{S} species per patch. Gene interaction networks might range from traits governed by additive genetic variance to different network topologies taking into account epistasis and pleiotropy to produce a trait distribution with different variance for each species in each patch (Figure 2)[19, 20, 21, 23, 24, 41, 42].

EQUATION 1: FITNESS IN MULTITRAIT GRAPHS

EQUATION 2: TRAIT VALUE OFFSPRING

Trait distributions obtained from additive or non-additive processes are then used to obtain each interaction strength extending previous food web models [43, 44, 45]. We generalize the function γ_{ixy}^t , represented as a Gaussian function describing the rate with which predator (or competitor) y with trait value z_y^t consumes prey (or competitor) x with trait value z_x^t in patch i at time t , as

$$\gamma_{ixy}^t = \frac{1}{N_\alpha} \left(\exp \left[- (z_y^t - z_x^t)^2 \right] + 2\alpha \left[\text{sgn}(z_y^t - z_x^t) \left(1 - \exp \left(- (z_y^t - z_x^t)^2 \right) \right) + \text{sgn}(\alpha) \right] \right), \quad (1)$$

where N_α is a normalization constant, $\text{sgn}(\mathcal{X})$ is the sign function and α is the interaction selection asymmetry. For $\alpha = 0, -1$, and 1 , predators or competitors prefer common prey or competitors, rare prey or competitor with more distant trait values, and rare prey or competitors with less distant trait values, respectively. The interaction strength, a_{ixy}^t , between species x for a specific intraspecific niche width ($ianw$) of the species y in patch i at time t can then be approximated as

$$a_{ixy}^t = \int_{ianw} \gamma_{ixy}^t D(x)^t D(y)^t dx dy, \quad (2)$$

where $D(x)$ and $D(y)$ are the density of the two species, respectively.

The community matrix containing the interaction coefficients between species x and y in patch i at time t and the connectivity obtained from the species interspecific niche width ($ienw$)[43, 44] is given by $\mathcal{A} = [a_{ixy}^t]$. The phenotypes after interaction selection for each prey or competition selection asymmetry scenario and before reproduction can be used to calculate fitness using a fitness gradient approach in the additive scenario [46] or without having to assume a particular fitness function in the non-additive scenario [47]. Fitness will then determine the ecological dynamics that is represented as a spatial network of local interaction networks.

We will run the model for many generations with each iteration containing interaction selection, mating and migration to compute the community matrix and the Jacobian for a gradient of dispersal values, following the dispersal between patches i and j , using the dispersal matrix, $\mathcal{D} = [d_{ij}]$. We will use the Jacobian to obtain the S-map or other stability methods to study the effect of gene interaction networks, interaction selection asymmetry, intra- and inter-specific niche width, and dispersal dynamics on the stability of hierarchical networks in the meta-ecosystem [12, 48].

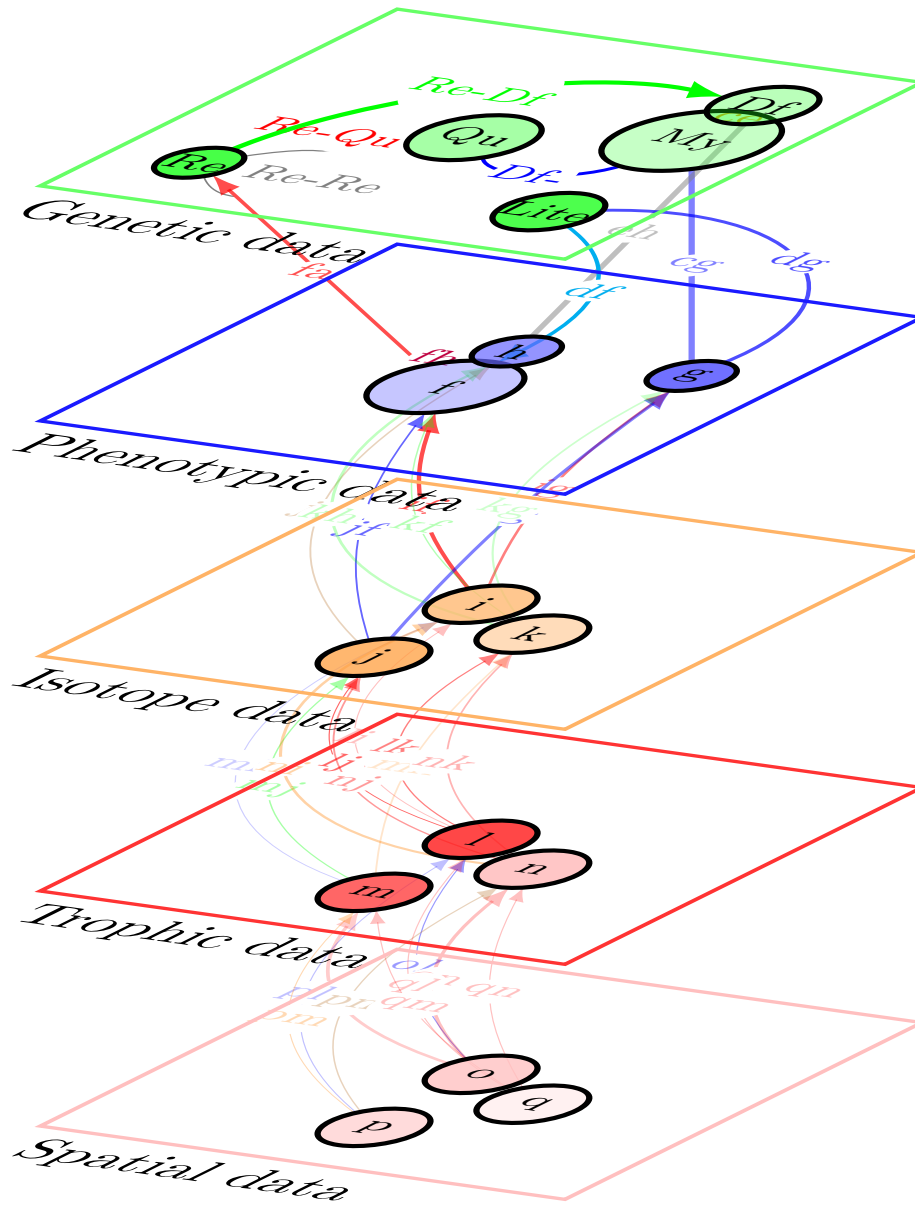


Figure 2: Example of the Threespine Stickleback multidimensional individual based database containing genetic, phenotypic, isotope, trophic and spatial data for around 1k individuals. This data is essentially QTL mapping (genetic data), morphological metrics of traits (phenotypic data), isotope and diet data and the spatial location (i.e., habitat type data).

4.2 Data sources

We will combine animations, data mining for pattern inference (Box 1) and process-based modeling (Box 2) to tangle genetic, phenotypic, ecological and spatial data. The following are the available datasets:

Projet Lac database Fish communities from the majority of lakes of Switzerland Webpage: <https://www.eawag.ch/en/departement/fishec/projects/projet-lac/> The database contains approx. 60k fish individuals, 40k morphometrics and 8k sampling actions each containing spatial coordinates, morphological traits, species abundances, habitats and DNA data.

Whitefish community database Whitefish community from the majority of lakes of North Alps. The database contains more than 2k fish individuals. Sampling locations, morphological, ecological and microsatellite data is available in the

Dryad database <https://doi.org/10.5061/dryad.k183ft7>

Threespine Stickleback database The Threespine stickleback fish database contains genetic, isotope, diet, morphometric, ecological and spatial data for around 1k individuals. Hosting these datasets in the SDSC platform might allow us to 1. Featuring animations of the Fish collection part of the FishEc database in a future permanent exhibition hosted by the Natural History Museum in Bern. The animations will highlight the interactions from genes to landscapes integrating genetic, phenotypic, ecological and spatial data (see M1 and M2 below). Please check <https://www.nmbe.ch/de/unser-angebot/unser-angebot/fuer-forschende/wirbeltiere> section Ichthyologie with Projet Lac being one of the projects contributing the most to the current FishEc database and <https://www.nmbe.ch/de/museum/aktuelles/monster-im-saft> for the status of this permanent exhibition. 2. Inferring patterns of interdependence throughout new uploads of source data (M3 and Box 1). 3. Analyzing process-based models to explore biodiversity dynamics scenarios with the aim to infer the processes that best reproduce the empirical patterns observed in the databases (M4 and Box 2).

4.3 Work Packages, Milestones and Deliverables

This project will strengthen the feedback between data-scientists and the community of scientists working with networks in ecology and evolution. This fusion will produce two methods packages and two scientific papers. Below we describe the timeline describing the milestones, the deliverables and the timing to release the packages in public repositories (Table 1).

CREATE TABLE LATEX

Work packages WP1: Data visualization (Lead: SDSC, contribution data: GENOM, ECOSYS, and BIOD) WP2: Web spidering for source data and data integration (Lead: SDSC, contribution: MAIN and BIOD) WP3: Data mining (Lead: SDSC and MAIN, contribution: CSYS) WP4: Process-based modeling (Lead: SDSC, MAIN and CSYS)

FIGURE MILESTONES

Milestones - M1.1: Uploading source data to Renga Platform - M1.2: Visualization and

animations of the source data as interdependent networks. - M2.1: Web spidering to assist in the search of databases of the missing source data - M2.2: Animation package of the Fish collection part of the FishEc database in the permanent exhibition hosted by the Natural History Museum in Bern - M3.1: Implementation of methods for inferring patterns in interdependent networks (Box 1) - M3.2: Efficient code implementation for inferring the interdependence among networks - M4.1: Implementing, debugging, running and analyzing the meta-ecosystem eco- evolutionary network modeling scenarios (Box 2) - M4.2: Analyzing the meta-ecosystem eco-evolutionary network modeling scenarios parametrized using the inference patterns obtained in M3

Partner WP Year and Quarter P 2019 L P P 1 2 SDSC SDSC GENOM MAIN 3 SDSC MAIN CSYS M3.1 4 MAIN SDSC CSYS M4.1 1 ECOSYS BIOD M1.1 BIOD 2 M2.1 3 M1.2 4 M2.2 1 2020 2 3 4 D1 M3.2 D2 D3 M4.2 D4

Table 1

Deliverables D1: Data visualization package to gain insights of the interactions across biological networks from genes to metaecosystems. The package will be used as highlights in the permanent exhibition of the natural history museum in Bern. We will produce a public github repository and a reproducible research document in Jupyter and Renga. D2: Data mining and inference patterns and processes in interdependent networks We will produce the inference package in interdependent networks. The package will be uploaded and maintained in a github public repository. Together with the inference package, we will produce a reproducible research document in Jupyter and Renga. D3: Scientific paper focusing on inference patterns in interdependent networks We will aim to a top general scientific journal. A reproducible research document will be developed in github, Jupyter and Renga. D4: Scientific paper focusing on process-based methods in interdependent networks We will aim to a top general scientific journal. A reproducible research document will be developed in github, Jupyter and Renga.

5. Requested Resources

5.1 Staff

5.1.1 Data science expertise

A two years data scientist from the SDSC will be required to guarantee the accomplishment of the word packages WP1 to WP4, and the deliverables D1 to D4 (see Requested resources and contributed resources below). The following are the key contributions for each stage of the project:

1. WP1: Software development to complement/improve the existing visualization and animations tools for multilayer networks. Many libraries are rapidly emerging to integrate, analyze, and visualize patterns in multilayer networks [123], yet add-hoc implementations in the existing tools or the development of new ones will be required to gain insights of the features of interact-

ing biological networks. 2. WP2: Machine Learning techniques and web spidering to assist in the search of databases to improve data integration among genetic, phenotypic, ecological and spatial data. 3. WP3: Efficient code implementation to overcome the computational challenges of inferring patterns of interactions connecting two or more networks across the webs of life (Figure 1 and Box 1). 4. WP4: Analysis and debugging of the process-based modeling scenarios taking advantage of the inference patterns obtained in WP3 to decipher the mechanisms explaining the observed patterns in the integrated datasets (Box 2). 5.2. Compute and storage resources 5.2.1. Summary We will need data storage and computing resources for running the visualizations, and the simulations using the data-driven inference of patterns and processes. The following are the working packages requiring computing and storage resources: CSWP1: Running the visualizations for a small dataset containing 5-10k individuals with genetic, phenotypic, ecological and spatial data. We would require to scale the visualization for this subset to a much larger dataset containing approximately 60k fish individuals, 40k morphometric measurements, and 8k sampling actions each containing spatial coordinates, morphological traits, abundances, habitats and DNA data. We expect the need of approximately 4 GPU during 6 months (see *SDSC_full_proposal_requested_resources_Melian2018.xlsx*) to run the visualizations and the animations using the

CSWP3: Implementing Gibbs sampling to fit the interaction coefficients for the interactions within and between the genetic, phenotypic, ecological and spatial data (Box 1). Our estimated need for CPU (cores) is of 35 for a 8 months period. We expect to be working with the following data size: 1) Data containing 5-10k individuals each containing a 20x20 gene interaction matrix and a 40x40 phenotypic interaction matrix. The spatial data contains 0.1k-0.5k sites each containing between 1 and 10 habitats. Each individual has coordinates within a habitat. 2) Data containing 50k individuals each containing a 40x40 gene interaction matrix and a 80x80 phenotypic interaction matrix. The spatial data contains 0.5k-1k sites each containing between 1 and 10 habitats. CSWP4: Implementing methods to fit process-based models to the empirical patterns (Box 2). We will be working with matrices as described in CSWP3.

5.2.2 Software packages

The following is the list of packages and skills to accomplish the deliverables D1 to D4: 1. Main skill: Spark GraphX to combine visualization, exploratory analysis and computation of metrics to infer network patterns crossing two or more networks. Additional skills: Implementation of packages in python (pymnet), java (gephi), and/or octave/julia (muxviz) languages. 2. Main skill: Javascript/Jquery to enrich the data sources proposed to be analyzed with external databases. Additional skills: Implementation of codes in python, Ruby or others to search databases. 3. Main skill: TensorFlow. We have been working at a very preliminary stage with a julia wrapper for TensorFlow 4 and we are open to learn from other packages to implement the methods required in this proposal to make them more efficient and robust. Additional skills: Hidden random Markov fields models or Bayesian methods to efficiently compute and/or reconstruct interaction networks using genetic, phenotypic, ecological and spatial data.

6. Contributed resources

Dr. Victor Eguiluz, team CSYS, will join as a senior scientist during the spring 2019 to work in the work packages WP3 (Data mining and inference) and WP4 (Process-based modeling). The main task of Dr. Eguiluz will focus on (Table 2): 1. Team up with the scientific staff of the SDSC to gain visualization insights of the ecological and evolutionary processes underlying the data sources provided represented as multilayer networks. 2. Implementing, debugging, running and analyzing the proposed meta-ecosystem eco- evolutionary network modeling scenarios (Boxes 1 and 2). 3. Implementing linear algebra techniques to contrast the stability-diversity properties of multilayer networks between scenarios driven by the empirical patterns and theoretical scenarios using null or neutral models. 4. Drafting scientific paper focusing on inference patterns in interdependent networks (Deliverables D2 and D3) Table 1 shows the contributed resources from the staff. Please see the *SDSC_{full_proposal_requested_resources_Melian2018.xlsx}document for a view of the tasks based in inference packages (Deliverables D1 and D2 with SDSC and work packages WP3 and WP4 with SDSC based methods in interdependent networks (deliverable D4) Supervising accuracy of the genetic data included based methods in interdependent networks (deliverables D3 and D4) Supervising accuracy of the genetic and based methods in interdependent networks (deliverables D3 and D4)*

References

- [1] C. Darwin. *On the Origin of Species*. Harvard Univ. Press, Cambridge, 1964.
- [2] D. J. Futuyma and M. Slatkin. *Coevolution*. Sinauer Associates Inc. Sunderland, Mass., 1983.
- [3] J. N. Thompson. *Relentless Evolution*. University of Chicago Press, Chicago, 2013.
- [4] R. MacArthur. Fluctuations of animal populations and a measure of community stability. *Ecology*, 36:533–536, 1955.
- [5] R. M. May. *Stability and Complexity in Model Ecosystems*. Princeton Univ. Press, Princeton, USA., 1973.
- [6] C. Jacquet, C. Moritz, L. Morissette, P. Legagneux, F. Massol, P. Archambault, and D. Gravel. No complexity-stability relationship in empirical ecosystems. *Nature Communications*, 7:12573, 2016.
- [7] J. A. Dunne, U. Brose, R. J. Williams, and N. D. Martinez. Modeling food-web dynamics: complexity-stability implications. *Aquatic food webs: an ecosystem approach*, Andrea Belgirano, Ursula M. Scharler, Jennifer Dunne, and Robert E. Ulanowicz (eds), pages 117–129, 2005.
- [8] E. T. Thebault and C. Fontaine. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329:853–856, 2010.
- [9] S. Allesina and S. Tang. Stability criteria for complex ecosystems. *Nature*, 483:205–208, 2012.
- [10] Johnson S., Domínguez-García V., Donetti L., and Muñoz M. A. Trophic coherence determines food-web stability. *Proceedings of the National Academy of Sciences of the U. S. A.*, 111:17923–17928, 2014.
- [11] A. Mougi and M. Kondoh. Food-web complexity, meta-community complexity and community stability. *Scientific Reports*, 6:24478, 2016.
- [12] D. Gravel, F. Massol, and M. A. Leibold. Stability and complexity in model meta-ecosystems. *Nature Communications*, 7:12457, 2016.
- [13] P. Alberch. From genes to phenotype: dynamical systems and evolvability. *Genetica*, 84: 5–11, 1991.
- [14] J. S. Arnold. Constraints on phenotypic evolution. *The American Naturalist*, 140:S85–S107, 1992.
- [15] V. Debat and P. David. Mapping phenotypes: canalization, plasticity and developmental stability. *Trends in Ecology and Evolution*, 16:555–561, 2001.
- [16] A. Wagner. *Robustness and evolvability in living systems*. (Princeton University Press), Princeton, USA. 2005.
- [17] W. G. Hill, M. E. Goddard, and P. M. Visscher. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics*, 4:e1000008, 2008.
- [18] S. K. G. Forsberg, Sadhu M. J. Bloom, J. S., L. Kruglyak, and Ö Carlborg. Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nature Genetics*, 49:497–503, 2017.

- [19] F. W. Stearns. One hundred years of pleiotropy: A retrospective. *Genetics*, 186:767–773, 2010.
- [20] A. Eyre-Walker. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences*, 107:1752–1756, 2010.
- [21] G.P. Wagner and J. Zhang. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics*, 12:204–213, 2011.
- [22] T. F. C. Mackay. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nature Reviews Genetics*, 15:22–33, 2014.
- [23] L. North, T. and M. A. Beaumont. Complex trait architecture: the pleiotropic model revisited. *Scientific Reports*, 5:9351, 2015.
- [24] M. Pavličev and J. Cheverud. Constraints evolve: Context dependency of gene effects allows evolution of pleiotropy. *Annual Review of Ecology Evolution and Systematics*, 46: 413–434, 2015.
- [25] T. G. Whitham, J. K. Bailey, J. A. Schweitzer, S. M. Shuster, R. K. Bangert, C. J. LeRoy, E. V. Lonsdorf, G. J. Allan, S. P. DiFazio, B. M. Potts, D. G. Fischer, C. A. Gehring, r. L. Lindroth, J. C. Marks, S. C. Hart, Wimp G. M., and S. C. Wooley. A framework for community and ecosystem genetics: from genes to ecosystems. *Nature reviews Genetics*, 7:510–523, 2006.
- [26] N. Loeuille. Influence of evolution on the stability of ecological communities. *Ecology Letters*, 13:1536–1545, 2010.
- [27] C. Fontaine, P. R. Guimarães, S. Kéfi, N. Loeuille, J. Memmott, W. H. van Der Putten, F. J. F. van Veen, and E. Thébaud. The ecological and evolutionary implications of merging different types of networks. *Ecology Letters*, 14:1170–1181, 2011.
- [28] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of Complex Networks*, 3:203–271, 2014.
- [29] Y-Y. Ahn, J. P. Bagrow, and S. Lehmann. Link communities reveal multiscale complexity in networks. *Nature*, 466:761–765, 2010.
- [30] M. De Domenico, M. A. Porter, and A. Arenas. Muxviz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks*, 3:159–176, 2014.
- [31] E. S. Morrison and A. V. Badyaev. The landscape of evolution: Reconciling structural and dynamical properties of metabolic networks in adaptive diversifications. *Integrative and Comparative Biology*, 56:235–246, 2016.
- [32] C. M. Schneider, A. A. Moreira, J. S. Jr. Andrade, S. Havlin, and H. J. Herrmann. Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences of the U. S. A.*, 108:3838–3841, 2011.
- [33] M. Kivela. Multilayer networks library. http://www.plexmath.eu/?page_id=327, 2013.
- [34] M. Cantor, M. M. Pires, F. M. D. Marquitti, R. L. G. Raimundo, E. Sebastián-González, P. P. Coltri, I. Perez, D. R. Barneche, D. Y. C. Brandt, K. Nunes, F. G. Daura-Jorge, S. R. Floeter, and P. R. Guimaraes Jr. Nestedness across biological scales. *PLoS One*, pages 1993–2002, 2017.

- [35] U. Bastolla, M. A. Fortuna, Ferrera A. Pascual-García, A., B. Luque, and J. Bascompte. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458:1018–1020, 2009.
- [36] P. P. A. Staniczenko, J. C. Kopp, and S. Allesina. The ghost of nestedness in ecological networks. *Nature Communications*, 4:1391, 2013.
- [37] K. Mueller. *Construction and behavior of Multinomial Markov random field models*. Graduate Theses and Dissertations. 11419, Iowa State University, USA, <http://lib.dr.iastate.edu/etd/11419>, 2010.
- [38] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B (Methodological)*, 36:192–236, 1974.
- [39] D. J. Harris. Inferring species interactions from co-occurrence data with markov networks. *Ecology*, 97:3308–3314, 2016.
- [40] N. J. Clark, K. Wells, and O. Lindberg. Unravelling changing interspecific interactions across environmental gradients using markov random fields. *Ecology* (doi:110.1002/ecy.2221), 2018.
- [41] N. Solovieff, C. Cotsapas, P. H. Lee, S. M. Purcell, and J. W. Smoller. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, 14:483–495, 2013.
- [42] D. Melo and G. Marroig. Directional selection can drive the evolution of modularity in complex traits. *Proceedings of the National Academy of Sciences of the U. S. A.*, 112:470–475, 2015.
- [43] N. Loeuille and M. Loreau. Evolutionary emergence of size-structured food webs. *Proceedings of the National Academy of Sciences of the U. S. A.*, 102:5761–5766, 2005.
- [44] K. T. Allhoff and B. Drossel. When do evolutionary food web models generate complex networks? *Journal of Theoretical Biology*, 334:122–129, 2013.
- [45] C. J. Melián, F. Baldó, B. Matthews, C. Vilas, E. González-Ortegón, P. Drake, and R. J. Williams. Individual trait variation and diversity in food webs. *Advances in Ecological Research*, 50:207–241, 2014.
- [46] P. R. Guimarães, Jr., M. M. Pires, P. Jordano, J. Bascompte, and J. N. Thompson. Indirect effects drive coevolution in mutualistic networks. *Nature*, 550:511–514, 2017.
- [47] J. P. DeLong and J. P. Gibert. Gillespie eco-evolutionary models (GEMs) reveal the role of heritable trait variation in eco-evolutionary dynamics. *Ecology and Evolution*. doi:10.1002/ece3.1959, 2016.
- [48] E. R. Deyle, R. M. May, S. B. Munch, and G. Sugihara. Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B Biological Sciences*, 283:20152258, 2017.